A Bold Investment in Research for the University of Illinois

A proposal to
the University of Illinois at Urbana-Champaign

FINAL v2.0

Executive Summary

In order for the University of Illinois at Urbana-Champaign to attain its vision of being the pre-eminent public research university in an era where research across all disciplines is critically dependent upon digital resources and services, the university must coherently accelerate the evolution of its digital research infrastructure. As a leading institution in computing and information technology with a culture of innovation, many efforts have arisen over the years that have been the product of this culture. While the largely organic development of digital infrastructure, resources, and support services has served campus researchers well, the lack of a coherent and unified approach is an impediment to this campus remaining at the forefront of research and limits its opportunity to be the leader in digitally-enabled research across all disciplines. In order to continue to excel, this university must make a bold move to proactively develop the coordinated digital support ecosystem required to realize its vision of pre-eminence in research.

Campus-wide efforts have established largely independent services that have brought value to the research community. However, these fall short of achieving the depth and breadth of potential value, and miss opportunities to promote research excellence that would be provided by a coherent research IT support ecosystem that is connected to the national and international research IT infrastructure. The Year of Cyberinfrastructure looked carefully at a broad cross section of the research community on the campus and revealed the lack of coherence and equity of access to resources, services and support by researchers. The Year of Cyberinfrastructure Final Report[1] provides a set of recommendations based on the findings of the team. These recommendations also align strongly with the Illinois Strategic Plan[2] which relays a charge to be agile and to lead change. Prioritizing a comprehensive and cohesive research IT support structure positions Illinois to successfully respond to the charge put forth by campus executive leadership.

In October of 2015, the Deputy CIO for Research IT position was created in the CIO's office to focus on the research IT needs of the campus with the specific charge to provide leadership and strategy to deliver cost-effective IT services and support to colleges and departments in support of the research mission on the Urbana-Champaign campus. This proposal is the expression of the vision developed: a dynamic research support environment in which a broad portfolio of resources, services, and support are easily discoverable and accessible to the campus research community. This has been driven by the Illinois Strategic Plan, the Year of Cyberinfrastructure recommendations, the experience of the team of more than 60 participants directly involved in this proposal to date, and the inputs received in response to a draft of the proposal from the IT Faculty Advisory Committee, the Academic Senate IT Subcommittee, the Engineering IT Research Working Group, and the many faculty and staff who responded during the open comment period provided to allow broad input to the proposal draft.

This proposal builds upon investments by the campus in the past several years to establish campus-wide services such as the CARNE ScienceDMZ, the Illinois Campus Cluster Program, and the Research Data Service by creating a research support environment that encompasses the much broader needs of our campus research community. These efforts, coordinated by the Deputy CIO for Research IT and often in

[1] https://www.ideals.illinois.edu/bitstream/handle/2142/88444/YearofCI_FinalReport.pdf and provided in this document as Appendix I due to frequent references to this document

[2] http://strategicplan.illinois.edu/

collaboration with efforts in various units around the campus, will provide a coherent set of resources and services that will bring value to our researchers and allow them to focus more upon their research and less upon the supporting technologies and services. The result will enable our researchers to be more competitive in the grant proposal process and to excel in their work, while making Illinois more attractive for faculty recruitment and addressing some facets of faculty retention.

The proposed activities that fall into four broad categories are, briefly, as follows:

- **Research IT Support**—one of the primary themes that emerged from the Year of Cyberinfrastructure
    - **Research IT Portal:** will create a nexus where researchers and research support professionals efficiently and effectively find, access, and use critical resources, services, and support available to them from across the campus.
    - **Research User Services:** will provide a seamless, integrated experience for *all* researchers to identify the resources available to support their work and gain access to support necessary to make use of those resources.
    - **Research Applications and Software Development Support:** will provide access to deeper expertise in the areas of supporting diverse research applications and research software development.
    - **Data Visualization and Analysis:** will provide access to deeper expertise in the analysis and visualization of a broad range of data types.
    - **Communications and Marketing:** will provide a channel through which to raise awareness of the available resources and services on campus and to develop the stories of the successes of our research community.
- **Research Computing Services**—a portfolio of flexible computing resources enabling the necessary agility to be competitive and to meet the wide range of research needs on campus.
    - **CARNE (ScienceDMZ):** an established network ecosystem designed to facilitate research activities ranging from wide area collaborations to local innovation on network protocols.
    - **Illinois Campus Cluster Program:** an established computing resource that supports tightly-coupled parallel computing jobs that must execute within a particular site, across low-latency interconnects.
    - **High Throughput Computing Service:** a computing resource that will unlock existing idle computers to execute large numbers of independent or loosely-coupled tasks, provided to researchers at negligible cost.
    - **Virtualization in Support of Research:** a service that will provide managed, elastic computing resources and easily customized software stacks to meet researcher-specific needs, and that migrates and scales seamlessly.

    - **Cloud Computing in Support of Research:** an emerging service providing expert support that enables researchers to quickly and easily leverage highly elastic and scalable commercial solutions.

- **Data Support**—the most frequently called for need during the Year of Cyberinfrastructure
    - **Research Data Service:** an established service that advises on data management planning and implementation, provides data management training, and develops and operates the Illinois Data Bank, a repository for public access to Illinois research data.
    - **Sensitive Data Services:** a service providing a campus-supported, compliant solution that allows researchers to store, process, and transmit sensitive or high-risk data in acceptable and risk-mitigated ways.
- **Research IT Strategy & Planning**—supports the evolution of the research IT environment through needs identification and assessment to support data-informed planning and decision-making.
    - **Needs Collection and Analysis:** will conduct ongoing needs assessments to track the evolving IT needs of the campus research community.
    - **Cyberinfrastructure Master Plan:** codifies a vision for the development of the campus' ecosystem of digital technologies and services in support of research.
    - **UIUC IT Architecture:** facilitates the integration of research activities into the infrastructure and services of the campus by creating a technical definition and description of IT infrastructure, platforms, and services, how they interrelate with one another, and the interfaces and access methods available for use.

These efforts will be complemented by support and coordination from the Office of the Deputy CIO for Research IT. This will include business and technical processes to enable rapid, hassle-free provisioning of computing resources and development of the financial models and financial management to support these initiatives along with the critical administrative and project management support.

Though Illinois faces significant financial challenges, it is time to make a bold investment to plot a course for the university that allows it to emerge as the destination for faculty, postdocs, graduate students, undergraduate students, and research staff to participate in world-class research in the modern digital age.

Introduction

In an era where research across all disciplines is critically dependent upon digital resources and services, the University of Illinois at Urbana-Champaign must coherently accelerate the evolution of its digital infrastructure in order to attain its vision of being the pre-eminent public research university. The University of Illinois is a technology-rich environment and a leading institution in computing and information technology with a culture of innovation. Many efforts have arisen over the years that have been the product of this culture. While the largely organic development of digital infrastructure, resources, and support services has served campus researchers well, the lack of a coherent and unified approach is now an impediment to this campus remaining at the forefront of research and limits its

opportunity to be the leader in digital technology-enabled research across all disciplines. Strategic plans and oversight committees have provided excellent guidance, but have largely been passive activities. In order to excel, this university must make a bold move to proactively develop the digital support ecosystem required to realize its vision of pre-eminence in research.

While the prevailing practice of researchers assembling needed resources and services has allowed those with the ambition, or more frequently the absolute need, to advance their fields, this has primarily only brought benefit to those researchers and their collaborators. Providers of resources and services have brought value to the research process, but this has been done in a largely disjointed manner that has also tended to favor the more advanced users of technology. The research University of the 21st century must have a fundamental foundation of research IT infrastructure and services—available to all—to properly support modern research practices and to be competitive and preeminent in the academic community.

> *"Growing investment in cyberinfrastructure is crucial to accelerating scientific discovery and engineering innovation across all disciplines."*
>
> Jim Kurose, Assistant Director, NSF CISE
> "NSF Fiscal Year 2017 Budget Request"

Some level of coherence has been brought to this ecosystem only by the efforts of those participating in multiple campus research IT initiatives. Illinois is now taking a more significant step forward in rationalizing the enterprise IT environment under the leadership of CIO Mark Henderson. In partnership with Vice Chancellor for Research Peter Schiffer, there have been efforts to provision and support IT infrastructure in support of research. However, Illinois is falling behind its peers in support of research by not bringing this same coherence to the research IT of the university. Campus-wide efforts have established largely independent services that have brought value to the research community but fall short of the depth and breadth of value possible, and the ability to promote research excellence, that would be provided by a coherent research support IT ecosystem that is connected to the national/international research IT infrastructure.

The efforts of the Year of Cyberinfrastructure[3], as encapsulated in the Year of Cyberinfrastructure Final Report[4], have made it clear that the research support landscape of the campus not only lacks coherence but is also very uneven across the academic units and sorely needs a vision for IT in support of research, and the ability to realize and evolve that vision in light of the evolving needs of the research community and the technologies available to implement that vision. The Year of Cyberinfrastructure covered a broad cross-section of the campus research community with participation from 12 of the 14 colleges and involving more than 200 participants across 27 focus group discussion, complemented by an additional 183 researchers responding via an online survey. The Year of Cyberinfrastructure Final Report provides a series of initial recommendations to address the needs of our researchers. Our university requires a dynamic research support environment in which a broad portfolio of resources, services, and support are easily discoverable and accessible to the campus research community.

[3] http://cyberinfrastructure.illinois.edu/

[4] https://www.ideals.illinois.edu/bitstream/handle/2142/88444/YearofCI_FinalReport.pdf and provided in this document as Appendix I due to frequent references to this document

The Illinois Strategic Plan[5] illustrates a changing landscape coupled with the challenges research institutions face, and the necessary steps and investments that need to be made in response. Illinois has a rich history of scientific discovery and the corresponding public outreach that improves the quality of life. The Illinois Strategic Plan continues with a charge to be agile and to lead change rather than being managed by it. Building a comprehensive and cohesive research IT support structure positions Illinois to successfully respond to the charge put forth by campus executive leadership.

The world of grant-based research is becoming increasingly competitive with decreasing success rates for proposals submitted to agencies and foundations across the board. The success of our researchers in these competitions is vital to their success and critical to achieving the research mission of the university. The proposed environment will enable greater competitiveness of our researchers' proposals resulting in higher success rates. Further, features of the proposed environment will allow some research proposals to be submitted that were not possible in the past (e.g. those requiring a HIPAA environment). These will lead to increased research funding to the university that will generate additional overhead and, in some cases, directly fund these services via cost-recovery methods.

Many peer institutions have recognized the need for establishing research IT resources, services, and support, and this is becoming a differentiating factor in the recruitment and retention of top research faculty. Though many factors come into play, it has become clear that a comprehensive and cohesive research IT support environment is not only a differentiating characteristic of a world class institution, it has become an expectation of top faculty.

THE WHITE HOUSE

Megan Smith, US Chief Technology Officer, presented "Harnessing the Power of Data, Technology and Innovation to Unlock Talent...and Unlock the Possible". The number of federally sponsored and privately subsidized initiatives to build infrastructure, develop innovative ways to leverage it, and grow talent is going to increase because the need for the output continues to grow.

National priorities that are White House initiatives include Big Data R&D; National Strategic Computing Initiative; Manufacturing, Robotics, and Smart Systems; Understanding the Brain, Smart Cities, and Computer Science Education. National Science Foundation (NSF) CISE (Computer and Information Science and Engineering) research has requested $995,000,000 to fund projects to address the national priorities and to continue priorities, like cybersecurity.

National Strategic Computing Initiative (NSCI) created to ensure the Unites States continues to lead in the development and deployment of cutting-edge computing systems.

"By strategically investing now, we can prepare for increasing computing demands and emerging technological challenges, building the foundation for sustained U.S. leadership for decades to come, while also expanding the role of high-performance computing to address the pressing challenges faced across many sectors." (https://www.whitehouse.gov/blog/2015/07/29/advancing-us-leadership-high-performance-computing)

[5] http://strategicplan.illinois.edu/

Though Illinois faces significant financial challenges, it is time to make a bold investment to help plot a course for the university that allows it to emerge as the destination for faculty, postdocs, graduate students, undergraduate students and research staff to participate in world class research in the modern digital age.

Creating a Modern Digital Infrastructure in Support of Research

Using lessons from the Year of Cyberinfrastructure, we propose to build upon the investments of the past several years in establishing campus-wide services such as the CARNE ScienceDMZ, the Illinois Campus Cluster Program, and the Research Data Service in creating an environment that encompasses the much broader needs of our campus research community and provides the resources and services along with the critical support necessary to enable our researchers to excel. This effort must:

- maintain a clear understanding of the evolving needs of the campus research community for research support IT;
- create and evolve a vision and roadmap for research IT for the campus that is responsive to its needs, accommodates the rapidly evolving technologies available, and appropriately interfaces the campus to the national and international ecosystems;
- bring coherence to the campus research IT infrastructure through leadership in campus-wide initiatives that coordinate the research IT efforts of campus units, and by implementing the vision defined by the Cyberinfrastructure (CI) Master Plan;
- assure that the research IT ecosystem and the enterprise IT environment complement each other in a manner that allows innovations in the research IT space to be developed and migrated, when appropriate, to the enterprise IT environment and vice-versa;
- ensure that robust and effective communication strategies are in place such that researchers and those who support them can readily identify the IT resources and support services available to them;
- incorporate the work of campus researchers to develop specific capabilities into the campus research IT infrastructure thereby broadening access to those capabilities.

The work of the Office of the DCIO for Research IT will occur within the context of the Illinois Strategic Plan in order to support the vision, mission, and strategic goals of the university.

> *"Advanced cyberinfrastructure accelerates [the] pace of discovery and innovation in all areas of inquiry and enables [a] platform on which cross- and inter-disciplinary research thrives."*
>
> Jim Kurose, Assistant Director, NSF CISE
> "NSF Fiscal Year 2017 Budget Request"

All of these activities combined will facilitate the Office of the DCIO for Research IT's effort to coordinate research IT activities within the units. The Office of the DCIO for Research IT will provide direct oversight for some campus-wide research IT initiatives such as the Research IT Portal, Research User Services, Illinois Campus Cluster Program, ScienceDMZ, and Cyberinfrastructure Master Plan in addition to others. It will also coordinate and foster coherence among these activities in conjunction with other research IT efforts within campus units (e.g. University

Library, NCSA, IGB, Beckman Institute, Technology Services, LAS, Engineering, etc.), leveraging the substantial experience and expertise of the participating units to bring additional value to the entire campus. Importantly, the Research IT activities will leverage the IT Power Plant—a collaborative partnership between campus IT organizations to provide IT services and support in a more consistent, reliable, and cost-efficient manner.[6] The Office of the DCIO for Research IT will also be a resource to faculty, research professionals, students, technology professionals, and units in furthering the research IT needs of the campus. It will serve as a trusted hub of information for finding services, resources, partners, and collaborators, a function facilitated by its information gathering and dissemination activities.

The Office of the DCIO for Research IT will also provide support functions for research IT initiatives. This will include business and technical processes to enable rapid, hassle-free provisioning of technology resources along with developing the financial models and financial management to support these initiatives. This will be done in collaboration with the Technology Services Business and Finance office and be complemented by administrative and project management support. Recognizing that not all faculty will be in a position to provide funds for services supported on a cost recovery basis, services of this type will include an amount of resources available generally to the campus research community that will be managed through an allocations process (described below). Finally, in addition to ongoing collection and analysis of the needs of campus researchers, the Office of the DCIO for Research IT will establish an advisory committee to help guide the development of resources, services, and support; it is anticipated that this will leverage other such advisory activities of the campus.

The Office of the DCIO for Research IT will be the proponent not only of projects that it leads but also of projects that it coordinates with other campus units. Further, this office will provide a means of effectively and proactively coordinating efforts to tap into state and federal sources of funds to move our campus forward in providing a unified and coherent digital research ecosystem.

Elements of a Campus Research IT Ecosystem

To build upon recent efforts toward realizing a dynamic research support environment, and to utilize the findings and recommendations of the Year of Cyberinfrastructure, we provide a discussion of the initial elements of a campus research IT ecosystem that will enable research excellence at the University of Illinois.

> *"Stop building new services if you're not going to support us in using them, provide us training in using them, or make sure we actually know they exist."*
>
> Year of Cyberinfrastructure survey respondent

Research IT Support

Although there is a clear and pressing need to foster existing efforts and establish new resources and services on our campus, this will not be effective without the proper support being made available in order to utilize those resources and services. For some researchers, appropriate support is all that is needed. For these researchers, the resources and services exist but they are either unaware of them or have no support in maximizing the benefits they can gain from existing opportunities. Broadly, we categorize these efforts as Research IT Support, which has several components.

[6] http://itpowerplant.illinois.edu/

Research IT Portal

The Research IT Portal will be a nexus for researchers and those IT professionals providing research support to efficiently and effectively find, access, and use critical resources, services, and support available to them from across the campus. The Research IT Portal will be a source of online information about these resources and services and, as a key component of the campus research service model, the portal will be an integral point of contact with the research community and a critical strategic resource for building sustainable relationships with researchers.

As identified in the Illinois Strategic Plan, it is imperative to develop innovative strategies to continually improve the research process by making it easier for campus to meet key goals. The Year of Cyberinfrastructure Final Report emphasizes that a single point of access to IT resources is needed. Access to these resources is an important facilitating condition for successful research and, in many cases, determines the critical path of the research process. The report describes focus group comments expressing frustration about the service experience, including, considerable time spent searching for resources, uncertainty about the process to acquire tools, time lost searching for resources to learn how to use software, and difficulty finding skilled staff and matching projects with staff availability. Additional survey data shows that only 63% of respondents were aware of Blue Waters, 56% were aware of Box, 32% ICCP, 17% XSEDE, 16% RDS, and 5% CARNE, from a list of 18 technology services (including an "other" option). It will also be imperative to maximize the number of researchers aware of the resources. An anonymous quote from the Year of Cyberinfrastructure survey stated: "Stop building new services if you're not going to support us in using them, provide us training in using them, or make sure we actually know they exist." The portal will help to reduce the time it takes to go from formulating a research question and designing the research effort to searching, acquiring, and using research IT resources.

The Research IT Portal will provide a comprehensive and coherent listing of existing campus resources and services that may be leveraged to help researchers. The Research IT Portal will be more than the technology on which it is based. It will be the foundation of a campus-wide sociotechnical system providing a high quality service. For example, imagine a solution that integrates a database of campus IT resources with the University Library search portal. In this scenario, the search results could be complemented with a new window entitled "Campus Resources." The University Library site has a frequently used "Ask a Librarian" chat service which receives 500 chats per week mid-semester and maintains very positive user-feedback. This is powered by well-trained Library staff and a customer service tracking and management system called IWonder.[7] Building on the Libraries' "Ask a Librarian" chat service, we will add a corps of IT resource subject matter experts to answer questions and, as needed, direct the researcher to other research services. We will develop a strategy for staffing this service in consultation with the service owners that we will partner with for this project.

The Research IT Portal will leverage various existing campus resources and the development team will assess the feasibility of using these resources. The portal will utilize service to drive the backend of the portal. This includes the integrating The IT Power Plant (e.g. Concierge and Help Desk services), University Library Portal, and unit level Help Desks will be used in a new portal-driven, coherent, and federated support model. The portal will interface with extensible campus resources including data and

[7] https://chat.library.illinois.edu/

knowledge bases with IT resource availability, teaching and learning resources, IT professional and research staff expertise, software product pages, user communities, and so on.

Here are some identified resources that will be utilized to develop the portal:

- The eXtreme Science and Engineering Discovery Environment (XSEDE)[8] User Portal[9] will be used as a model for the Research IT Portal.
- The portal will leverage knowledge and experience from a series of research studies on IT portals.
- The portal will integrate available training resources.
- The portal team will collaborate closely with the IT Communications and Marketing.

Agile methods[10] will be used to plan, design, develop, deploy, and support the Research IT Portal. The startup phase will begin with planning (months 1-3), during which a systems analysis study and stakeholder analysis study will be conducted. In conjunction with efforts of the Communications and Marketing team, we will develop a new centralized repository of information and determine the feasibility of extending existing resources. The team will evaluate licensing agreements, work-flow processes, organizational constraints, governance issues, human resource constraints, and technology limits. These efforts will lead into design (months 4-6) and development (months 7-10) phases culminating in the first production deployment of the Research IT Portal (months 11-12). The Research IT Portal will then transition to an operational mode with ongoing feature additions and enhancements informed by the work of the Needs Collection and Analysis team and feedback from the campus research community.

The startup phase will require a project lead (1 FTE), a user-centered design specialist (1 FTE), web developers (2 FTE), and technical writers and content developers (2 FTE). After the startup phase, 1 FTE will be needed on an ongoing basis to deliver feature additions and enhancements and to integrate additional resources and services as they are identified or established, supported by web developers (2 x 0.5 FTEs) to maintain portal software. A 0.5 FTE will be needed on an ongoing basis to keep the content and the knowledge base fresh as well as provide training and support the web developers. The Research IT Portal team will coordinate with nearly all other research IT teams to assure that the resources and services they offer are properly incorporated into the portal. Further, IT Power Plant services will be leveraged, in particular coordination with the Help Desk for support service integration and Web Hosting services for deploying portal software (e.g. Liferay[11]) to support the portal environment by a 0.5 FTE resource. The portal environment will include an enterprise portal with an integrated content management system supporting mobile technologies and enterprise file synchronization. This phase will require $30,000 annually for the portal software and hosting. The team has identified Liferay as a platform. However, as the team gathers more user requirements, it will evaluate different platform options for the portal infrastructure.

[8] http://www.xsede.org
[9] http://portal.xsede.org
[10] http://agilemethodology.org/
[11] https://www.liferay.com/

Research User Services

Many researchers struggle to navigate the complexity of our highly-fragmented IT service landscape, often unable to answer a simple “where do I get assistance with technology X?” question. The Year of Cyberinfrastructure Final Report indicated a significant lack of awareness of research IT resources in the form of both services and staff and a diversity of IT skills among researchers. One researcher commented, “[We] can often find a collaborator but not necessarily the tech support."

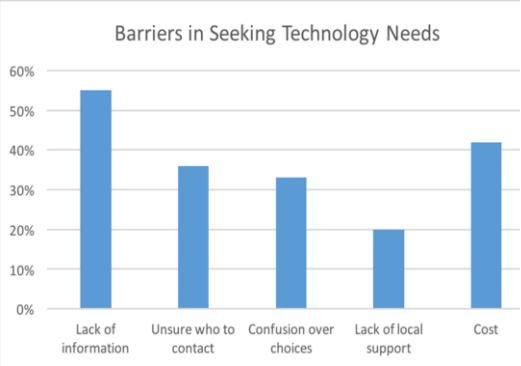

Figure 1: Barriers in identifying technology solutions to support research needs as identified during the Year of Cyberinfrastructure.

The Research User Services initiative will provide a seamless, integrated experience for all researchers regardless of unit or services needed by building upon existing support structures that will improve the researcher experience. By improving documentation efforts, creating a concierge facilitator service and research IT consulting office, and leveraging a professional campus help desk, we will:

- simplify the identification and request of IT services,
- improve response and resolution times,
- resolve more issues via self-service or consistent 24-hour help desk support.

Improving documentation of research IT services is critical to researcher success and the effectiveness of research support efforts. The Year of Cyberinfrastructure Final Report identifies this need and notes that many researchers prefer self-service solutions. Documentation must include what is available, where it’s available, and how best to use it. It must also be provided via the Research IT Portal in formats understandable by both researchers and the IT staff who support them. Documentation is critical to the success of all levels of research user services, including the campus help desk, concierge facilitator service and the research IT consulting office. The Research IT consulting office will be responsible for facilitating the creation and maintenance of relevant documentation, with assistance from communications and marketing staff. The documentation effort will leverage the University’s Knowledge Base, already in use by AITS, Technology Services, ATLAS, College of Education, DIA, UIS, and the UIC School of Public Health.

Documentation alone will not meet the IT needs of researchers. Fifty-three percent of all Year of Cyberinfrastructure respondents prefer a hands-on approach, working collaboratively with technology experts. The help desk, concierge facilitator service, and research IT consulting office will provide this multi-level hands-on support to researchers.

The campus Help Desk, targeted to be a 24x7x365 service via the IT Power Plant, will provide a first-point-of-entry human interface to researchers. It will be the resource to which researchers can turn when they can’t quickly identify a solution via the Research IT Portal and may be able to resolve many basic issues of researchers.

> *“[We] can often find a collaborator but not necessarily the tech support."*
>
> Year of Cyberinfrastructure Focus Group Researcher

For example, one researcher asked, “When we have datasets larger than will fit in one Excel file, what do you do?”

The Help Desk will leverage robust documentation services to efficiently identify available resources or connect researchers to a concierge facilitator or staff in the research IT consulting office. The Help Desk is expected to be available to all units and funded centrally, so no additional Research IT staff are required.

The concierge facilitator service will work to provide support in matching the needs of resources with existing resources and services on campus. The College of Engineering has employed research technology facilitators since 2011 and launched a similar service in 2014, and other IT professionals serve in smaller-scale, informal facilitation efforts. The roles are inspired by the example of the XSEDE Campus Champions, feedback from the Year of Cyberinfrastructure, and CIO Mark Henderson’s call for a research concierge service.

The Research IT concierge facilitator service will:

- Consult with researchers on what technologies to explore and integrate.
- Build awareness of available research IT resources, focused on access and integration of those resources into their research programs.
- Foster community and coordination among unit-level technology facilitators.

While some units may have local staff performing similar roles, the campus concierge facilitator staff will support this need across all units and provide campus-wide leadership in this area. A team of 3 concierge facilitators will be necessary.

The Research IT consulting office will provide more advanced support to researchers as needs become more complex and specific. Consulting office staff will be well-versed in all research IT resources on campus and able to engage faculty in more in-depth interactions over a longer period of time. When appropriate, consulting office staff will escalate an issue to the appropriate specialty service, such as the Illinois Campus Cluster Program or the High Throughput Computing service. The proposed consulting office is an industry standard intermediate support environment common in research environments. One ACES researcher commented, “[We need] better tools or expertise to analyze data, even to know if data makes sense. [We have] lots of data created, but not enough knowledge/tools to know if it is any good."

A team of 10 FTE or more will be necessary to support the diversity of research IT services and needs on campus, built over time as new services are deployed. We will begin with 2 staff, growing to 6-8 FTE within 12 months, and 10+ FTE after 24 months. Approximately 0.7 FTE will be committed specifically to the ICCP, HCP, and CARNE efforts.

> *"[We need] better tools or expertise to analyze data, even to know if data makes sense. [We have] lots of data created, but not enough knowledge/tools to know if it is any good.”*
>
> Year of Cyberinfrastructure Focus Group Researcher

Research Training Services

For all technologies, new and old, there is a need to be trained on how to best use them. During the Year of Cyberinfrastructure focus group

discussions, researchers requested opportunities to receive training on a much broader range of tools and services than is available today. They also requested training options to a higher level of expertise than currently available. A member of an ACES Agricultural and Biological Engineering focus group lamented that "grad[uate] student labor is not as qualified, [there is] erosion in program development training and software development, [and a] lack of IT skills hamper[s] research." Researchers often do not know how to get started with technologies unfamiliar to them, and some technologies have long learning curves to develop sufficient expertise to be beneficial.

> *"Are Illinois graduate students prepared to meet the expectations of their fields in the 21st century?"*
>
> Year of Cyberinfrastructure Closing Symposium

Several activities are proposed to support training efforts across campus. First, consolidation of training events into a master calendar system will take place over the summer. An email list will be set up for trainers to communicate and promote events. During the fall of 2016, an expanded training resource encompassing all training events will be developed, and incorporated into the Research IT Portal. An affiliation to the Software Carpentry Foundation will enable key training resources and regular workshops. We will additionally seek to expand personnel support for broad-impact, general-interest training activities. In addition to the Software and Data Carpentry curricula which cover introductions to data science tools, training topics identified so far include:

- general interest subjects (e.g., high performance computing, Linux, and statistical packages
- project and workflow management
- domain libraries (Boost, GEANT4, etc.)
- quantum chemistry tools
- data analytics training
- digital humanities (text corpora analysis, linguistic data, etc.)

Portal content will incorporate high-quality on-demand resources (such as MOOCs or tutorials) as these are identified by trainers and the content coordinator for the portal.

A total of 6.95 FTE are requested for this service. Personnel support (exclusive of ongoing commitments) should include:

- 1 FTE content coordinator/service manager
- 2 FTE lecturers/instructors through Computational Science and Engineering to provide general technical training and support tools coursework offered through various departments (These funds will be used to enable domain experts to contribute directly to training and coursework development through overage or buying a portion of their time directly.)
- 4×0.5 FTE TAs (2 FTE annually) through CSE to support the same
- 0.25 FTE web developer to develop and support the training portal site
- 0.2 FTE IT training support for virtualization efforts
- 0.5 FTE graduate assistant to support general training coordination (beyond civil service admin. support)
- 1 FTE civil service administrative support

In addition, we request in monetary support:

- $10,000 annual commitment to Software Carpentry Foundation (affiliation)

Trainers and support personnel will seek to promote services to and engage IT professionals, faculty, and research staff as well as students. This serves two purposes: to make potential users of training aware of events, and to engage skilled practitioners as occasional instructors as interest permits. We anticipate that this service will address the Year of Cyberinfrastructure recommendations to provide training for researchers.

Research Applications and Software Development Support

A variety of campus researchers need deeper expertise in the areas of supporting diverse applications and research software development, as outlined in the Year of Cyberinfrastructure Final Report (specifically, Theme #1). That theme identifies the significant need for more access to expert personnel with advanced technical expertise. Researchers desire the ability to utilize staff with deeper technical skills for hours to years in order to perform specific tasks or participate in research projects requiring significant efforts in software development or research applications support.

As we consider support for research computing resources at Illinois beyond Help Desk or Consulting Office capabilities, researchers need either more advanced support or longer term assistance. This is also reflected in Recommendation #4 of the Year of Cyberinfrastructure Final Report which states, "...*provide researchers with access to professional staff who have specialized technical skills*." Many researchers need access to technical specialists to adapt the use of computation to research fields that have not traditionally considered computation as part of their toolkit, or to adapt research software to operate on computers supported at Illinois. Additionally, support for the use of domain application software, whether it be commercial or community supported, is often critical for the success of research efforts. Finally, research groups may need assistance developing software for particular research needs. The report identified needs with respect to interfacing with research instrumentation, software development, and support for use of third party software. Such a research support service represents a unique assembly of expertise that will participate in campus training development and delivery.

Provisioning such a specialized, deeply engaged capability requires consideration of the breadth and depth of expertise required to provide value to the Illinois community. The "rules of engagement" need to be developed to ensure that any such centrally provisioned effort is available on a fair and equitable basis. Mechanisms must be in place to allow for direct support from specialized staff to be involved for longer-term support of research efforts where it is needed and funded as part of an independent research grant.

Many of those with the need to access such experts typically only need that expertise for short durations and they cannot secure the funding to hire such a person full time for a minimum of two years. Even those with funding face challenges with the timing of Notice of Awards and lengthy hiring and onboarding processes. Thus "on-demand" access is a universal need. Mechanisms will be established to allow for fair and equitable distribution of staff resources funded by campus to support campus research. A model that has worked well for XSEDE is to make the support allocable in the same manner as compute and data resources: via a request and review process. We have the opportunity to apply these existing methodologies to provide application and software development support resources to researchers.

Principles embraced by XSEDE for allocation of staff resources include a) definition of a time limit for the support, b) availability of the support at any point in the research cycle, and c) requirement for the definition of a work plan (albeit high level) which is used to guide the work and to indicate criteria for the project to be considered complete. These principles help ensure availability of support to a research group at no cost. A key to making this approach successful is that support needs to be considered as a collaborative effort, with contributions by the research group as well as the supporting staff.

Two established campus groups providing application and software development support are NCSA and IGB. They provide a basis upon which to build a campus-wide capability covering the diverse needs of the campus research community. We propose using the rule-of-thumb from XSEDE that involvement in a project typically costs 25% of an individual for 6 months, thus each FTE will complete approximately 8 projects per year. Some exploratory software projects may require high initial effort over a longer duration, with subsequent effort being reduced.

Based on demands identified via the Year of Cyberinfrastructure efforts and through other channels such as the Illinois Campus Cluster and XSEDE user base, a core of 10 FTE is needed to enable both advanced applications support and research software development support. We propose that 5 FTEs of each type of support be made available to the campus community and allocation of that staff time be done as described above. Additionally, as the projects look to adopt agile methods, being able to tap into agile project managers to assist with project planning on a temporary basis would be enormously valuable. This provides a base of staff that can be assigned to projects that provide funding in an agile and responsive way. The appropriateness of this level of support will be assessed in conjunction with the Needs Collection and Analysis team as they conduct their outreach across the campus.

We will also consider alternative methods of providing application support for research and educational computing to complement the established methods. These would include using complex graphical user interfaces such as provided by Engineering IT services. Cost and licensing restrictions have motivated Engineering IT to provide access to the applications to any students or faculty on the Illinois campus. Taking this further, we propose to investigate hybrid models, including combination of interactive front-end software combined with batch backend computing to fully meet the need in instruction and research.

As part of a leading-edge effort to provide the highest quality applications and software development support, we need to provide state of the art software engineering capabilities to the campus research community. This includes software code repositories and issue tracking software. Besides the provisioning of these capabilities, we need to provide and encourage training of best practices for the professional staff and research community to work productively together. This would result in more productive workflow and improved collaboration between research teams.

Collectively, Research Application and Software Development Support is requesting new support for 7.5 FTE, combined with requests for 2.5 FTE from the Illinois Campus Cluster Program and High Throughput Computing service requests, for a total request of 10.0 FTE. Additionally, this service depends on support from project management and program management, and requests $91,000 startup costs for application delivery and software engineering capabilities, with an ongoing maintenance expense of $33,000 annually.

Data Visualization and Analysis

44% want better access to data analysis
46% want better access to data visualization
Year of Cyberinfrastructure Final Report

Visualization and data analysis have become ubiquitous in the campus environment. Common software tools utilized at various levels from the student population to researchers and even non-technical staff likely contain elements of visualization and/or data analysis. This makes the field an ideal candidate to benefit from any number of infrastructure improvements and those most critical are generally applicable campus-wide. There are, however, associated difficulties. Foremost, a rigorous framework is needed to properly narrow the scope of the field as currently the exact delineation from other related fields is unclear. Defining the scope will enable us to differentiate areas that need growth from established areas that simply need coordination or promotion, e.g. GIS services. Beyond that, a set of currently unavailable resources have been identified that would benefit this large community. These pertain to software and staff as well as capability.

From the Year of Cyberinfrastructure Final Report, "researchers were presented with a list of seventeen technology resources expected to be most important to a respondent's research in the next few years. Their top three most commonly selected resources are all data-related. These resources are: storage services (35.8%), analysis services (26.9%), and visualization services (22.0%)." See Figure 2 as well, where both data visualization and analysis were in the top-4 challenges discovered during the Year of Cyberinfrastructure. Improved or additional visualization and data analysis resources will benefit a large portion of the campus community, including researchers aware of these capabilities and new or emerging disciplines that could enhance research outcomes through data visualization and analysis. These resources are typically related to software (training, support, user communities) and/or staff expertise (consulting, advanced support, collaborative efforts). As such, the success of a resource may ultimately be determined by the significance of its user base and/or the growth rate of this population.

Challenge	% Respondents	# Respondents
Analyzing	67%	219
Acquiring	56%	182
Visualizing	41%	133
Storing	40%	132

Figure 2: Visualization and analysis were both in in the top four data challenges identified during the Year of Cyberinfrastructure.

Through evaluation of visualization and data analysis resources currently available to Illinois students, staff, and faculty, areas were identified where provisions are inadequate or nonexistent. What follows is a recommended series of actions based on that evaluation for the next 12 months to enable missing capabilities, manage improving resources, or carefully create the foundation for an optimal cyberinfrastructure. The first step is to solidify the organization of visualization and data analysis through community consensus, and we will map the visualization work already being done by engaging the Grainger Library, NCSA, Beckman, and others. The largest appropriate context centered on the visualization and data analysis community should be surveyed to reach consensus on how the field will be described.

The nature of visualization and data analysis makes for an ideal standard use case for the evaluation of infrastructure improvements. With visualization firmly integrated into the cyberinfrastructure in that capacity, attention will be turned to deploying new campus-wide services, namely a data analysis and visualization consulting group and funded support for enabling data analysis and visualization services

generally for the campus. We also see the need for better remote visualization capabilities and will therefore analyze current successful deployments, and develop a plan to standardize, incorporate, and support such capabilities in the future. Finally, we will formalize the responsibility to continually survey external resources to identify possible targets for deployment at Illinois.

Many of the above tasks are in early stages and we believe leaders for planning and deployment are available through volunteer efforts of experts here at Illinois. Deploying new services will require additional staff, but current models exist that alleviate costs at the university level. For a consulting group, minimal funding of managerial staff and researchers is required to deliver an effective set of student consultants, but these costs may be recovered through campus research teams engaging consultants. However, general visualization and data analysis support will require an investment of ongoing funds to support the experts who would see this work through. While these costs could similarly be recovered, ongoing funding provides two benefits: 1) we could start with a fraction of a single FTE which is likely absorbable by groups already performing similar work, and 2) when such is not possible, this offers flexibility to more easily fund new staff for these groups.

> *"Need Data Visualization tools: lots of data, getting more difficult to understand and even more difficult to present data in a meaningful way."*
>
> Year of Cyberinfrastructure Focus Group
> ACES Ag & Bio Engineering
> March 2, 2015

There are documented Returns on Investment[12,13] (ROI) for having staff dedicated to produce a collaborative Data Visualization and Analysis team at Illinois. This leads directly to increased collaboration of efforts and techniques across campus. Additionally, for our student population, we accelerate data skills and have potential to certify data analysis and visualization capabilities. The university may thus begin to feed into the loop of educating, retaining, and attracting top researchers and graduate students. These skills also apply in the efficient management of the university facilitating discovery of relationships and patterns between business and operational activities as well as quickly identifying emerging trends. We therefore propose that an initial 5 FTEs be made available to the campus community as described above in the Research Applications Support and Research Software Development Support section. A comparison with the XSEDE environment substantiates this as consistent in scale and pragmatic.

Communications and Marketing

The need for research-focused communication was a common theme in the data gathered during the Year of Cyberinfrastructure. A survey respondent attending a Year of Cyberinfrastructure faculty focus group meeting shares, “Mostly as a new professor, I have no idea what is out there most of the time, and what information does reach me is largely irrelevant. It would be nice if the information put out there was more tailored to what people in specific fields would likely need, instead of a general large broadcast of everything that is bound to get lost in the shuffle.”

Staff from the Media, Information, and Communications (MIC) team in Technology Services will work with other existing departmental communications teams to map the communications landscape at the

[12] http://www.analyticbridge.com/profiles/blogs/4-business-benefits-of-data-visualization
[13] http://data-informed.com/top-5-business-benefits-using-data-visualization/

> *"Mostly as a new professor, I have no idea what is out there most of the time, and what information does reach me is largely irrelevant. It would be nice if the information put out there was more tailored to what people in specific fields would likely need, instead of a general large broadcast of everything that is bound to get lost in the shuffle."*
>
> Researcher
> Year of Cyberinfrastructure Focus Group

University of Illinois. The focus groups that are created through the Research IT Needs Collection and Analysis effort will determine what communications methods will best reach researchers. These two efforts combined will provide strong improvements in our ability to ensure researchers understand what resources are available, when new resources are developed, and how to find them in order to get started.

Many of the necessary communications and marketing tools are already in place and will be leveraged. These include targeted mailings, social media, electronic newsletters, digital signage, mailing list services, and video hosting. All can be leveraged for any number of campus communications efforts with no additional costs. The IT Power Plant is addressing the need for providing web hosting solutions, including several tiers of options which will give communicators and researchers access to tools that meet a range of service needs. With the Research IT Portal being the central hub that we want to drive researchers to, we will minimize hosting and service costs and be able to focus on content and marketing.

For development of communications content, one FTE will create traditional written content highlighting success stories where research IT resources were used. One FTE would create rich media versions (video, animations, etc.) of the best success stories to bring them to life for a wider audience. Two FTE would work on gathering data and creating a marketing strategy to best communicate with researchers. A 0.5 FTE from the MIC team will be dedicated to developing marketing and communications strategies and materials as part of the startup phase of the Research IT Portal and the Cyberinfrastructure Master Plan.

The same half of one (0.5) FTE would work with the Research IT Portal on continually promoting the service as well as curating content, while another half (0.5) FTE would work with colleges and departments on orientations and events for researchers.

> *"Make it clear what the campus even offers. Advertise! Reach out!"*
>
> Ludmila Sergeevna Mainzer
> Year of Cyberinfrastructure Closing Symposium
> "Storage and Computing Needs in Life Sciences"

Research Computing Services

Maintaining an edge in the increasingly competitive research world requires us to be ready to aggressively pursue any opportunity at a moment's notice. A rich cyberinfrastructure that provides a variety of flexible computing resources enables the necessary agility to be competitive and to meet the wide range of research needs on campus. Further, we must provide the support services necessary to guide researchers in using these resources, as they discover they need them, and rapidly develop customized solutions to research problems.

Creating a portfolio of research computing services begins with a number of foundational tasks. It will be necessary to: (1) conduct focus groups and gather additional input from researchers to more clearly understand their immediate research needs, informing the focus of our initial efforts; (2) streamline business and technical processes that obstruct the rapid provisioning of resources; (3) align existing, emerging, and proposed computing services to ensure that they address the requirements and constraints imposed by research needs; and (4) facilitate adoption of these services by providing small but useful allocations to researchers that allow them to explore and evaluate these computing resources.

The University of Illinois' total research and development expenditures in FY14 were $622M.
https://research.illinois.edu/sites/research.illinois.edu/files/upload/ovcr_researchreport_2015_100915.pdf

CARNE (ScienceDMZ[14])

The Campus Advanced Research Network Environment (CARNE) is a network ecosystem designed to facilitate research activities ranging from wide area collaborations to local innovation related to network protocols. Modeled after ESnet's ScienceDMZ, where the network is optimized for science data transfers, CARNE also encompasses the necessary tools for providing end-to-end support for such transfers. As highlighted by a Year of Cyberinfrastructure focus group participant, networks and the associated tools and support specifically optimized for research usage will alleviate current problems such as moving 1-2 TBs of data to other campuses.

CARNE, which consists of high-speed, low-latency, friction-free network paths for research support, is a necessary asset when seeking high-profile wide area collaborations and when competing for grant funds. CARNE serves as an onramp to strategic nationwide research network connections like Internet2, providing improved access to other national and international science resources such as national and international cyberinfrastructures, the Large Synoptic Survey Telescope project, and the Large Hadron Collider at CERN.

In its current state, however, CARNE is only a partial solution. Network infrastructure must be improved in buildings where bandwidth-intensive instrumentation and other sources and sinks for data exist. To maximize the return on expanding infrastructure, performance monitoring and performance tuning tools such as perfSONAR devices and a mesh of high-performance data transfer nodes are integral components for enabling end-to-end data movement to researchers. Dedicated FTE are needed to build platforms yielding a proactive support model coupled with self-service provisioning that enables the research community to get maximal benefit from the network. As tools and functionality increase, so must the development of local expertise through partnerships with the Research IT Support teams to directly support researchers. When discussing an efficient file sharing tool, Dr. Liudmila Sergeenva Mainzer, a Senior Research Scientist at NCSA and a Research Assistant Professor at IGB, supports the necessary outreach and education components. "Globus client is very easy to install and use, but people simply do not know about its existence and benefits." Dr. Mainzer goes on to suggest, "Education about the best use of appropriate technology is key."

[14] http://fasterdata.es.net/science-dmz

Recognizing that ScienceDMZ's were largely conceptual among higher education institutions, the National Science Foundation (NSF) responded with funding opportunities distributed through the Directorate for Computer and Information Science and Engineering to build and expand the utilization of ScienceDMZs. Since 2012, the NSF has released annual solicitations ranging from $500,000 to $5,000,000, which could potentially be future funding opportunities for Research IT Services, and specifically CARNE.

The 2013 IT Strategic Plan[15] highlighted needs and a path toward Illinois using IT to build competitive advantages and streamlining the user experience. These points are supported by the Year of Cyberinfrastructure Final Report. Expanding and enhancing CARNE with the necessary support tools will improve the technical performance of many research endeavors. Data-intensive research and expanding wide area collaborations coupled with pressure from granting agencies suggests that the usage of CARNE will grow rapidly.

Dedicated FTE and funding would provide a path for building the necessary support tools and developing research-centric support expertise, which will empower researchers to focus on their science without regard to the underlying enabling technology. Four (4) dedicated FTE, combined with an existing 0.25 FTE, with non-FTE expenses of $350K initially, scaled back to $300K annually, will work across the Research IT Services framework to develop CARNE. For the initial expansion phase, efforts will be conveyed via 0.05 FTE from the Research User Support Communications and Marketing group. On an ongoing basis, 0.20 FTE and 0.05 FTE from Research User Support and Training, respectively, will be leveraged.

Illinois Campus Cluster Program

Campus researchers increasingly need access to computational resources to conduct their research. High-performance computing (HPC) applications tend to focus on tightly-coupled parallel computing jobs that must execute within a particular site, across low-latency interconnects. Traditional HPC applications include computational fluid dynamics, quantum chemistry, and astrophysical simulations, while fields such as computational finance, digital humanities, and archaeology are relative newcomers. The Illinois Campus Cluster Program (ICCP) was formed to address this need and provide a professionally-run resource that researchers can invest in, regardless of disciple or departmental affiliation. Dr. Maria L. Chu, an assistant professor in Agricultural and Biological Engineering, recently had her research highlighted on the Illinois website.[16] "Running a model without a supercomputer would take me the rest of my life," she says.

> *"Running a model without a supercomputer would take me the rest of my life."*
>
> Dr. Maria Chu

As an established resource of over five years, the ICCP is operated as a high-quality, production service for all of campus upon which researchers can rely to conduct their research, fostering discovery and innovation. As a campus-level resource, the ICCP provides substantial cost savings to campus and to researchers while offering a path for those retiring existing compute resources to migrate and centralize their operations while still having a voice in the direction and governance of

[15] http://itstrategicplan.illinois.edu
[16] http://techservices.illinois.edu/news/2016/future-water

the resource[17], thereby making us better stewards of available funds. A key example of this is the Computational Science and Engineering (CSE) investment in the ICCP, made as an alternative to replacing their aging Turing compute resource. Both of these benefits are part of the Illinois Strategic Plan. Furthermore, access to the ICCP provides a competitive advantage to our researchers over other universities that do not offer such a resource. The ICCP has always leveraged value and expertise on campus, from tapping NCSA talent for designing and operating these systems, to helping drive the development of CARNE, a relationship that has greatly benefited the ICCP. More information about the ICCP can be found on its website.[18]

The ICCP is taking lessons learned from its first five years, with the recommendations from the Year of Cyberinfrastructure Final Report, and will modify its operational strategy to provide new ways for investors to gain access to and utilize the resource. The traditional hardware investment model will be augmented by the sale of compute cycles for as-needed use. This will allow for both node purchases and usage-based billing, meeting a broader set of researcher needs. Further, the ICCP is currently transitioning from building discrete clusters within the environment, to continuous deployment model that enables the program to adopt new technology for our investors at a faster pace. This will be done while maintaining the single, integrated environment for all campus cluster resources and minimizing investment risks.

As a resource, the ICCP requires many technical systems to ensure robust operations for all researcher investments, such as servers for management and operations, high-speed networking, and centralized high-performance storage. Most equipment providing these key functions are five years old and require a hardware refresh due to age. This proposal includes both a one-time request for $960,000 to refresh this equipment, and an annual hardware budget of $304,680 to maintain it continually. Additionally, the compute pods must be deployed as investor demand necessitates. A compute pod[19] with necessary 10-Gigabit networking equipment costs $179,000, and up to two of these could be required per year. The costs for the high-speed InfiniBand network will be passed on to only the investors needing such exceptionally high-performance communications.

As a service, the ICCP requires 8.55 FTE of staff time to design, configure, and operate equipment, resolve user issues, develop new features, and optimize systems performance. Current funding commitments account for 4.55 FTE, and we are therefore asking for an additional 4.00 FTE to augment the current team. The ICCP will also need 0.40 FTE to support the efforts of the Virtualization in Support of Research efforts.

The ICCP is most successful when it exists within the fully realized IT Research ecosystem, and its vision can only be realized with support from the Research IT areas of project management, Training, Research User Services, Research Applications and Software Development Support, Data Visualization and Analysis, Communications & Marketing, Needs Collections & Analysis, CARNE, and the Research IT Portal. Other proposed Research Computing Services areas complement the ICCP in ways that enable

[17] https://campuscluster.illinois.edu/governance/
[18] https://campuscluster.illinois.edu/
[19] A pod is a unit of growth within Data Center Shared Services. Each pod for the ICCP can hold 400-500 compute nodes.

broader workflows, such as the High Throughput Computing service, and the ICCP integrates with CARNE and shares its vision for moving data effortlessly to wherever researchers require.

High Throughput Computing

The High Throughput Computing (HTC) service proposes to create a new computational resource for Illinois researchers by tapping into the idle processors located in workstation labs across campus. This follows the model in use at the University of Wisconsin-Madison and many other research universities. The HTC service is targeted at the bulk execution of a large number independent or loosely-coupled tasks that can be individually scheduled on many different computing resources. The number of jobs that can be completed per unit time is typically of key importance. This is the case for large scale parameter searches, Monte Carlo simulations and some high energy physics analyses.

The College of Engineering alone has approximately 1,500 workstations in labs in that could be pooled together to create a significant HTC resource. Discussions with some stewards of those resources have expressed interest in participation in this effort. The researcher-facing front-end for this service will be operated in coordination with the ICCP, leveraging those staff and their engagement with campus research teams. A modest investment alongside the ICCP will see these idle compute cycles provided to Illinois researchers at little-to-no cost. Researchers will leverage the Research Data Services' Active Data Storage service as the data repository component of their HTC research efforts.

This HTC resource will enable researchers with the appropriate computing workloads, those who are just getting started, and in particular those who come from departments that are historically underfunded for technology, to access a very large computational resource to conduct their research. This makes them more competitive for grants and also provides a mechanism to transform the scale of computational work being done by some researchers. As a use-case, consider the Illinois researcher who can today only process their research on their office workstation. This researcher limits the scale at which they investigate their research, often without fully realizing it, whether by limiting the resolution at which they investigate the data or limiting the number of parameters they consider in their investigations. With the proposed HTC resource, that researcher could greatly improve the detail and accuracy of their research without having to use program budgets to acquire compute time. A resource that is available in this way can be offered to students of all levels, whether for instructional purposes or for compute time in support of Ph.D. research. As a participant in the recent Data Infrastructure Workshop suggests, "Cloud facilities on campus where unused nodes would be available for others to use. Everybody needs compute."

> *"Cloud facilities on campus where unused nodes would be available for others to use. Everybody needs compute."*
>
> Data Infrastructure Workshop participant

Creating such a resource is clearly aligned with the mission and goals set forth by the Illinois Strategic Plan. This resource will: fosterer scholarship, discovery and innovation; provide a transformative learning experience; and makes us better stewards of our current investments. In particular, it will unlock investments in largely idle resources on campus to enable research. The service will provide a competitive research advantage, improve student outcomes, and, by tying in with the ICCP, will streamline the user experience.

In the pilot phase, we will assemble a small team to evaluate HTC platforms alongside friendly research teams and workstation labs eager to partner with us. Service prototypes will be created and evaluated, settling on a final platform and project plan for how the solution will be rolled out on campus. After 12 months in pilot, the final service offering will be put into place and offered to campus with support and documentation. As the production service gains traction, the HTC team will approach stewards of other workstation labs to discuss participation in the HTC team's effort to capture otherwise wasted lab compute cycles. Also, research teams on campus who currently operate small (e.g. 20 workstation) HTC pools will be approached, in the hopes that they will willingly pool their resources with ours to make a larger and more cohesive solution for all of campus. Integration with national computing grids, such as the Open Science Grid, and commercial clouds, such as Amazon Web Services, will also be investigated.

Fortunately, the largest cost associated with the HTC service is already funded: the thousands of lab workstations previously purchased by departments. To leverage that investment, a startup budget of $25,000 and 2 FTE of staff will be necessary to operate the resource and support researchers. Ongoing, we request $15,000 annually for hardware augmentation and refresh. In the pilot phase, the effort will be focused on conducting trials, evaluations, and outreach. After transition to production operations and the service stabilizes, the staff will remain at 2 FTE, however effort will split. One (1) FTE will remain within the HTC service for administration and feature development and the remaining effort will shift to the documentation, training, and support areas. Close ties with the ICCP make the HTC service an incremental addition to the compute portfolio and minimizes effort.

Virtualization in Support of Research

The research community requires managed, elastic computing resources that allow them to easily customize software stacks to meet their specific needs, and that can be easily migrated from small-scale, local systems to larger scale campus and national cloud resources. User surveys conducted for the Year of Cyberinfrastructure indicated that this is of particular interest to researchers in the area of distributed systems and networks. One participant responded "If a researcher wants to do research in cloud computing, networking, distributed systems, then more open virtual machines environments and OpenStack clusters-testbeds might need to become available." Virtual machine and container hosting services have emerged as the standard methods for delivering flexible computing resources. Providing for virtual machine and container hosting services paired with a solution architecture and research support efforts is a critical component to maintaining the University of Illinois' reputation as a world class research institution.

Marketing and training materials, and a self-service web portal—all developed in collaboration the Communications and Marketing, Training, and Research IT Portal efforts, respectively—will be developed to allow researchers to explore how they might apply these technologies. Solution Architects and System Engineers will develop and maintain a library of virtual machine images and containers to support common research activities, and assist and coordinate migration of those virtual machines between computing resources. Success stories and case studies will be documented and publicized to serve as exemplars of virtual machines and containers in research.

Initial activities will focus on developing the foundations of the service. The Needs Collection and Analysis group will conduct surveys and focus groups to ensure we understand the needs of the research community. Findings will inform the development and implementation of capabilities to support virtual machines and containers executing in the environment of the Illinois Campus Cluster,

NCSA's OpenStack service, and the IT Power Plant offerings from Technology Services on Virtualization and Amazon Web Services. This will include gathering and analyzing user information and use cases which will be used to guide and prioritize the development of customized virtual machine images and containers. The Training and Research User Support teams will also use this information to determine how best to serve researchers.

Once the foundation is built, we will build customized virtual machine images and containers designed to efficiently solve a variety of research problems. Success stories will be documented and published as case studies that can be used to guide future researchers. The Training and Communications and Marketing Teams will leverage this information to develop new training seminars and workshops, and to explain the benefits of this service to the research community. The Research IT Portal will serve as a useful tool for outreach activities, and those groups will coordinate their efforts to effectively and efficiently engage with research customers.

The core activities of this service will require a modest investment in staff resources. Initially, a team of five people working part time should be sufficient. Two 0.4 FTE solution architects will recommend solutions including computing resources, workflows, and procedures to accomplish researcher goals. Solution architects will also provide deep technical support for complicated virtual machine and container-related problems beyond the scope of front line research support services. Two 0.4 FTE system engineers will build and maintain virtual machine images and containers, and assist with migrations between computing resources, augmenting the efforts of other groups and service providers. Additional staffing resources will be leveraged in cooperation with other Research IT groups, IT organizations, and the IT Power Plant. Staff will be needed to augment resources for Technology Services' AWS services, IT Power Plant Virtualization, ICCP, NCSA's OpenStack environment, Research IT Portal, Needs Collection and Analysis, and Training. NCSA's OpenStack cluster will need additional resources to increase capacity to host virtual machines. Additional hosting servers costing $30,000 are needed initially. Funding from cost recovery will fuel future expansion and upgrades of this service.

Cloud Computing Service for Research

Access to cloud computing resources is becoming a common expectation for researchers. Not only are research groups interested in how new cloud tools and services can be applied to their research, but granting agencies like the National Institutes of Health are increasingly requiring public cloud components in successful grant proposals. As Clair Sullivan, Assistant Professor in Nuclear, Plasma, and Radiological Engineering, asserts, "If we cannot expand beyond bare metal, we are dead." As articulated in a Year of Cyberinfrastructure focus group meeting, one participant asserted, "Scientists are using the cloud for real science." Many researchers interested in cloud computing do not currently have local expertise they can rely on to help evaluate technologies and architect solutions. With expert support from the IT community, researchers will be able to quickly and easily leverage highly elastic and scalable commodity solutions.

Researcher groups have traditionally taken responsibility for purchasing servers, configuring operating systems, and maintaining the software needed for their projects. Cloud vendors offer high-level building blocks which provide much of that same functionality out of the box. Allowing faculty and students to avoid busywork and devote more time to high-value research presents a competitive advantage. Streamlining students' IT experiences and teaching courses on industry-leading cloud platforms enables

more effective instruction and cultivates the skills employers want. By building on extremely scalable systems, instructors can innovate without concern for technological barriers.

The Year of Cyberinfrastructure Final Report recommends that we should "make research-critical technologies available to the campus community in a manner that ensures a low barrier to access and use." The initial phase therefore consists of efforts necessary to make Amazon Web Services (AWS) available to researchers quickly while minimizing the training that will be required. The technical tasks are focused around quickly provisioning accounts for research projects, federating with campus authentication and authorization platforms, and establishing the business process to do simple rebilling of usage costs for projects. These efforts have been ongoing since the start of 2016 and announcements for open access to these tools for researchers is imminent. We will also arrange on-campus training opportunities, gather feedback in order to fine tune the local execution, and make consultation resources available for users seeking in-depth assistance using AWS.

Beyond the startup phase, focus will shift towards building foundations to extend the campus environment into AWS to make it easier for researchers and others to seamlessly use both AWS and campus resources. This will include providing mechanisms for integration with campus networks, extensions of authentication and directory services, and best practice documents and training. Access will also be expanded to include the entire campus, including administrators and IT support teams.

This effort will be closely tied to a variety of other aspects of this proposal and other ongoing activities. This includes Training, Research User Services, CARNE, the Research IT Portal, and potentially all of the Research Computing teams. Particular emphasis will be put on training and collaboration opportunities in order to create a high level of support for research use of cloud services. This effort will intersect with IT Power Plant efforts and will expand to include other cloud services such as Microsoft Azure and Google Cloud computing. We will also partner with on-campus research-driven private cloud initiatives to ensure that researchers are able to use the most appropriate and effective resources available.

Building a capable team to support cloud computing will initially require one dedicated FTE (in addition to an existing commitment of 4 FTE by Technology Services) for the first six months of effort. Once the core infrastructure is in place, three FTE will be required to coordinate training and support customers. An annual budget of $25,000 will pay for all the necessary AWS resources for service management, service development, and campus-wide hands-on training.

Data Support

Data is a core component of research, and data-related concerns quickly emerged as a theme during the Year of Cyberinfrastructure activities. Almost all participants in nearly all demographics across the disciplines expressed concerns both in surveys and focus groups.

> *"When we have datasets larger than will fit on one Excel file, what do you do?"*
>
> Year of Cyberinfrastructure
> Focus Group Researcher

Data needs are diverse and complex because while infrastructure in the hardware sense is required, needs do not stop there and in fact very quickly bleed over into training, compute, analysis, visualization, preservation, etc. Two specific Data Theme recommendations came from the Year of Cyberinfrastructure Final Report:

- Enhance and expand the scope of efforts of the Research Data Service to address the evolving needs of researchers with respect to data.
- Develop, deploy, and support solutions that can manage restricted access to sensitive data.

Research Data Service

The Research Data Service (RDS)[20] was established through a program proposal following several campus-wide conversations revolving around data management, stewardship, and storage including a Year of Data Stewardship held in 2011-2012. The Offices of the Provost and Vice Chancellor for Research provided ongoing funding for the RDS in 2013; staffing began in 2014 and was completed by mid-2015. Primary services offered directly through the RDS include advising on data management planning and implementation, providing data management training, and developing and operating a repository for public access to Illinois research data, the Illinois Data Bank (launching in May 2016). The RDS program also came with one-time funding for NCSA to offer mid-performance, high-volume storage, known as Active Data Storage (ADS), as a cost recovery service to be promoted through the RDS. RDS is based in the University Library where the program's backbone comes from the strong interaction with the staff of many relevant Library units (including the subject Libraries, Scholarly Commons, Preservation, Archives, and Library IT). The program also partners with campus units including GSLIS, OVCR, Technology Services, and NCSA to enable more ready adoption and use of current services (e.g. Box.com) or establish additional services (e.g. ADS). For more details on RDS activities, see the 1st Year Annual Report.[21]

In the nearly two years since the RDS was formally established, it now stands as a central, customer-facing service. Many of the issues that arose during the Year of Cyberinfrastructure focus groups and survey results are the same that RDS struggles to overcome, namely uneven resources or those that are partially or wholly unsupported. Through realization of this Research IT Investment, IT-related gaps that the RDS has encountered (see RDS NewSoft Document in Appendix II) can either be addressed directly or assessed more fully working in concert with the Research IT Strategy & Planning needs assessment function (see below). In addition to the IT-related expertise that the RDS offers around data publication and preservation, the Research IT initiative will be able to further leverage the customer-facing nature of the RDS as a vehicle for multi-directional information transfer.

For the purposes of this proposal, no additional resources are being requested at this time for the RDS. As noted above, in 2013 the Vice Chancellor for Research initiated ongoing funding for the RDS with an annual budget of $450,000. This supports 4 FTE plus visiting staff (e.g. postdocs and GAs) and any associated travel, infrastructure costs for the Illinois Data Bank (i.e. storage), memberships (e.g. DataCite, ORCID), and marketing. At the same time, startup funds in the amount of $471,000 were allocated to establish the Active Data Storage service as a fully cost-recovery storage service for campus researchers in need of affordable, mid-performance, high-volume storage (stewarded by NCSA). ADS usage models include: online backups, storage for the hosting of large data sets, operational data storage and scratch space for analysis, mid-term data storage for archival needs and more. The ADS offers scalable storage from a few TBs to over a PB of storage exported as a private file system for each project and available through multiple transfer protocols with good transfer rates.

[20] http://researchdataservice.illinois.edu/
[21] http://go.illinois.edu/RDS_IDEALS

What is critically important here is the commitment of the RDS to coordinate efforts across the research IT initiatives in recognition of the need for holistic support. As a central service for campus researchers, the RDS is ready to play a role as an important bridge that connects researcher needs and IT resources.

Sensitive Data Services

Various researchers across campus at times need to store, process, or transmit sensitive or high-risk data. The Sensitive Data Service will serve researchers by establishing a campus-supported, compliant solution encompassing programs, environments, support, and controls, allowing them to work with sensitive data in acceptable and risk-mitigated ways. In some cases, this will enable the pursuit of research funding that requires such a solution; the absence of such to date has prevented researchers from submitting proposals to a number of funding agencies, particularly the National Institutes of Health (NIH).

Illinois has existing and potential needs for managing sensitive and high-risk data under a large variety of legal standards and requirements, including the Family Educational Rights and Privacy Act (FERPA), Health Insurance Portability and Accountability Act (HIPAA), Personal Information Protection Act (PIPA), International Traffic in Arms Regulations (ITAR), and others. These regulations all have security requirements and risks associated with misuse, leakage, loss, or other improper disclosure. Providing an environment that addresses compliance in this area is one of the significant needs identified by the Year of Cyberinfrastructure Final Report and is a clear implication of the fourth goal of the Illinois Strategic Plan to steward our resources. Researchers, administration, and staff have all recognized that data breaches in higher education and industry have become commonplace, and also that these compliance

NSF

The National Science Foundation (NSF) is the funding source for 24% of all federally supported basic research conducted across all higher education institutions. CISE/ACI (Computer and Information Science and Engineering/Advanced Cyberinfrastructure) recognizes the need to address cyberinfrastructure challenges at the campus level. Solicitations like the most recent CC*DNI encompassed $500,000 to $5,000,000 funding opportunities for network infrastructure as well as for programs to stimulate cyberinfrastructure adoption and the resulting innovative usage. The recent NSF solicitation for Cybersecurity Innovation for Cyberinfrastructure is another example of potential funding sources up to $1,000,000.

Since 2010, UIUC has received the most NSF grant dollars of any higher education institution.

NSF ACI has proposed a 6.3% budget increase ($995M) in “trying budgetary times”.

CISE FY 2017 request is shaped by investments in core research, education, and infrastructure programs as well as critical investments in NSF cross-directorate priorities and programs.

Enable/incent science and CI community to address data governance, lifecycle issues, sustainability

From NSF Director Cordova's testimony to Congress who speaks of NSCI and D4SDA.

“As the CIF21 investment sunsets in FY 2017, NSF will develop a subsequent, focused set of activities aligned with the Administration's new **National Strategic Computing Initiative (NSCI)** ($33.20 million) in order to focus efforts on advancing the Nation's computational infrastructure for science and engineering research. The rich topic of "Big Data," encompassing data science, data assimilation, data management, data policy, community building, and workforce development, will remain a strategic focus under the new NSF Data for Scientific Discovery and Action (D4SDA) activity, which will span research and research infrastructure.”

requirements, related liabilities and possible penalties apply. They are requesting accommodations, services, and controls that satisfy their specific compliance needs while enabling them to conduct business, collaborate, and research unimpeded.[22]

This effort also serves to fulfill the Illinois Strategic Plan by using IT to build a competitive research advantage for all faculty, researchers, students, and graduates. The question of performing research with high-risk data but without the benefit of a controlled environment or service means it is either done unsafely, exposing the sponsoring institutions to risk and liability, or not done at all, which deprives an institution of opportunities. Building this service will allow campus researchers to take advantage of the opportunities at hand.

Recently, some work has begun as a collaborative effort between the Office of the Vice Chancellor for Research, Technology Services, and NCSA to develop an initial HIPAA environment. Ongoing work will require further policy development around additional types of sensitive data and expansion of the sensitive data environment to support the policy and technical requirements.

Security infrastructure cost is expected to be $100,000-$150,000 annually for AWS cloud, and also $100,000-$150,000 for an onsite solution. Either or both may be chosen. If onsite hosting is selected, an estimated additional $100,000 initial investment will be needed to establish an appropriate computing and storage environment within existing building and infrastructure environments. It is expected that the services will be provided via a cost recovery model covering operational costs.

Beyond computing and storage resources, security risk assessments required by HIPAA and other risk compliance requirements will be needed and is expected to be $50,000 annually per solution chosen. This will need to be complemented by third-party risk assessments conducted in the initial year and at least every three (3) years at an estimated cost of $50,000 per assessment.

Management of these configuration will require 4-6 FTE each to provide customer consultation and technical stewardship of the environments. Each will also utilize 1.75 existing FTE to provide documentation, business and systems administration support.

One application that is a candidate for inclusion in the service and should be considered for evaluation and possible acquisition in the future is REDCap[23] (Research Electronic Data Capture), a mature, secure, web application that provides a platform for users to build and manage databases and surveys. This application also provides a platform for data sharing and shared database access between investigators and also between institutions. REDCap is available free of charge with support through an active consortium. Should the Sensitive Data Service provide REDCap to its customers, the need would include a technical FTE (or team) for installing and maintaining the software, as well as an administrative FTE (or team) to provide day-to-day assistance to REDCap customers. Illinois' Interdisciplinary Health Sciences Initiative (IHSI) staff member Ruth Sosnoff has scoped the REDCap platform for Illinois and found that universities employ as little as 1.5 FTE and as many as 20 FTE to provide REDCap support on their campuses. With the forthcoming Carle Illinois College of Medicine, we anticipate REDCap would require 3 FTE initially, growing to 10 over the course of two years. With a sensitive data service and a REDCap

[22] Although classified data does have compliance requirements it is out of scope for the Sensitive Data Service.
[23] http://project-redcap.org/

instance in place, Illinois may have a springboard from which to engage with researchers to understand the full scope of sensitive data needs.

It is important to note that the costs identified here relate to initially establishing and maintaining the environment. There will be additional efforts necessary to run the program and train the sensitive data environment users and support staff on compliant methods for using the infrastructure. This has been incorporated into the budget requests of the Training and Research User Support teams also being proposed.

The costs for researchers to deploy compute and storage resources within the environment are planned to be largely supported on a cost-recovery basis. While it is common to include such costs in proposals to agencies when such data is part of the research proposed, it is recognized that certain unique instances may be subsidized in whole or in part by the service or by other sponsoring entities.

Research IT Strategy & Planning

The series of components described previously represent the different parts of a bold investment in creating a leading IT environment to support research in higher education. To create this environment in alignment with the strategic mission of the campus and the needs of faculty and researchers requires tightly coordinated efforts in planning, informed by regular needs assessments, and knowledge of the changes in the broader technology landscape. The area of Research IT Strategy and Planning will be responsible for three major efforts to insure this alignment: gathering and analyzing needs information, creating and maintaining a Campus Cyberinfrastructure Master Plan, and providing a coherent architecture for the components of the plan including the interaction between components and with other campus IT resources.

Needs Collection and Analysis

The Year of Cyberinfrastructure Final Report's first meta-recommendation is to, "Establish an ongoing process for engagement with the campus community to continue to solicit feedback that captures their evolving needs." According to the report, feedback from researchers across the campus is vital to creating a productive research environment. Such data is also vital for a Cyberinfrastructure Master Plan to define our future state and provides critical input, feedback, and guidance to all research IT initiatives.

Responsive to this meta-recommendation, the Office of the DCIO for Research IT will conduct ongoing needs assessments to keep abreast of the IT needs of the campus research community. This will include a permanent version of the Year of Cyberinfrastructure efforts as well as regular scanning of what is developing nationally--among grant funding agencies and at peer institutions. This information will be complemented by input received from the appropriate IT governance committees (e.g., the IT Faculty Advisory Committee and the Academic Senate IT Subcommittee). The Office of the DCIO for Research IT will produce regular reports on identified needs and new areas of opportunity to guide the directions of research IT services. All of this information will also be used to drive regular updates to the Campus Cyberinfrastructure Master Plan and the evolving UIUC IT Architecture. The needs assessments will also act as a source of information for those wishing to propose research IT projects and as a source of identifying potential faculty collaborators for those projects.

The capabilities created to ingest this information from the research community will also be utilized to solicit feedback. Faculty and other research staff will be asked to assess the quality of the resources and

services that are provided and react to the evolutionary direction described in the Campus Cyberinfrastructure Master Plan. The feedback loop provided through this process will ensure that both the quality of services are improving and the roadmap established by the master plan is aligned with the needs of the community and is being achieved.

The needs collection effort will consist of meeting researchers through focus groups and conducting surveys which will be designed in consultation with the Survey Research Lab.[24] Data will be analyzed and reported back to the Research IT service teams. Progress and key metrics will be shared through the research IT Portal. Initial deliverables are expected to be available approximately six months after start-up. We will leverage existing IT advisory committees to review continuing efforts to ensure process improvements and adapting to the changing research environments. A service for faculty benefits from faculty participation in implementation and representation in advisory committees. Faculty advisory representation should cross disciplines ensuring input from the Science, Technology, Engineering, and Math (STEM) and Humanities, Arts, and Social Sciences (HASS) communities.

Strategic outcomes to our campus from this effort include:

- An ongoing communication channel with researchers to provide a continuous feedback and promote a long-term relationship.
- A data set focused on researchers to be used by Research IT to improve service offerings.
- Dissemination of information concerning technology to researchers who may be unfamiliar with resources available to them.

The target population for data collection is researchers on the Urbana campus. We use the term "researcher" because it is broad. Our goal is to assist the faculty of our campus and their assembled teams of experts with diverse skillsets. Those teams consist of research scientists, graduate students, post doctorates, visiting scholars, and other professionals who conduct research side by side with faculty. We include all of these researchers as candidates for inquiry.

The team has identified skillsets and resources required to initiate and continue the data collection and analysis program initiated by the Year of Cyberinfrastructure efforts. We propose a total of 3 FTEs overall to cover the roles of Team Lead (1 FTE), data collection teams (6 x 0.20 FTE) and data analysis (0.80 FTE). Project management and administrative support will be provided by the broader IT Research effort. Startup costs of $40,000 are estimated for designing the data set and collection process. Ongoing costs such as software licensing, room rental etc. will require $20,000 annually. Data collection teams are expected to come from realignment of duties of existing IT professionals already on staff.

Systematically planning the focus groups, data collection, and data analysis will allow us to properly tap into units and utilize expertise across campus. Our hope is that by providing the opportunities for discourse to take place, we will be able to put people in touch with services as well as share data about new and no-longer existing gaps in Research IT.

Cyberinfrastructure Master Plan

A Cyberinfrastructure Master Plan for the campus will provide a vision specifically for the development of the campus ecosystem of digital technologies and services in support of research. This master plan

[24] http://www.srl.uic.edu/

will allow campus to play a leadership role in the national and international IT ecosystems critical to the support of the collaborative efforts in which so many of our researchers are engaged. The intent of that document is analogous to a traditional facilities master plan—to provide a framework for guiding the long-term growth and expansion of the campus digital ecosystem to support research within the fabric of the Urbana-Champaign campus in a way that builds upon both existing and future efforts in the development, deployment, and support of digital research support capabilities.

Stewarded by the Office of the Deputy CIO for Research IT, the Campus Cyberinfrastructure Master Plan will be developed as a living document, updated as appropriate in light of evolving needs and technology solutions. This will serve as the campus roadmap in developing the Illinois research IT ecosystem. The Office of the Deputy CIO for Research IT will work with faculty, researchers, and technology professionals to support research needs and technology development efforts across the campus through leadership, oversight, and collaboration.

Beyond the strategic benefits of the Campus Cyberinfrastructure Master Plan, it has many practical benefits. Some funding agencies are requiring documentation of campus cyberinfrastructure as a part of the proposal process. A well-defined master plan can provide ready access to this information when building proposals. The master plan can also serve as a recruiting tool for faculty and research staff to communicate the assets available on campus and the future roadmap for expansion of those assets. The same details that are useful in recruiting are also useful for communicating to existing faculty and staff what is available to support them, filling the current gap of understanding that exists on campus regarding the resources which are available. As a respondent from a Year of Cyberinfrastructure focus group survey offers, "The resources and potential for integration of advanced CI on this campus are fantastic."

> *"The resources and potential for integration of advanced CI on this campus are fantastic."*
>
> Respondent from a Year of Cyberinfrastructure Focus Group Survey

The Cyberinfrastructure Master Plan represents a new effort on the campus. Initial efforts, anticipated to take 12 months, will require 1.25 FTEs dedicated to the effort with an additional 8-10 campus volunteers participating to form a core committee supplemented by subject matter experts as needed. Analogous to a facilities master plan, a consultant with expertise in the area of planning and higher education will facilitate the collection of information, the coordination of campus resources, and aggregate the outcomes to build an initial Cyberinfrastructure Master Plan. This is estimated to cost approximately $100,000. Following the initial phase of creating the Cyberinfrastructure Master Plan, ongoing efforts to maintain and update the plan will require 0.70 FTEs of dedicated effort working closely with the Needs Collection and Analysis effort to gain feedback and the Communication and Marketing effort for outreach to the campus community.

UIUC IT Architecture

The creation of a campus IT Architecture will provide a technical definition and description of IT infrastructure, platforms, and services, a description of how they interrelate with one another, and the interfaces and access methods available for use. The goal is to structure an environment that will be usable for researchers who want to tie their projects into campus-provided services so they can focus on

their research instead of duplicating available IT infrastructure and services. This will improve researcher productivity and competitiveness while reducing overall spend on IT and IT-related services.

On campus today, there is no coherent architecture describing the IT environment, creating a frustrating experience for those who attempt to find information about and use common IT services. The IT Architecture will provide information at multiple levels of depth and detail, with an initial emphasis on research-related needs. The IT Architecture will also be developed in conjunction with the Cyberinfrastructure Master Plan efforts in order to provide forecasting and direction for the development and lifecycle of campus-wide IT services.

The plan for developing an IT Architecture is to leverage the XSEDE architecture model used to manage a variety of IT services across multiple institutions in the national cyberinfrastructure environment. The initial phase will be to establish a collaborative campus team to work with the IT Architect to develop a high-level description of the general components of our current IT environment on campus. This will become the framework for the documentation of the next level of detail, which will describe specific service areas and where they fit relative to one another. These Level 1-2 Decompositions will provide a solid overview of the campus IT environment and its major components. In the next phase, service-specific teams will work in conjunction with the IT Architecture teams to develop detailed, service-specific interfaces in a Level 3 Decomposition. This includes templates and a standard methodology for the architecture documents themselves, as well as a standard process for documenting the driving use cases.

The IT Architecture team will be a mix of staff and volunteers from various points on campus. Two FTE need to be allocated to the IT Architecture effort to establish standard methods, coordinate the various teams, shepherd information gathering for use case development, write, edit, and collaborate on documentation, and provide connection points with Research IT Strategy and Planning teams and the other Research IT areas. The staff will also need to be involved in other IT efforts on campus driven by the Office of the CIO and other campus IT initiatives. One additional FTE of effort will be divided across multiple people who will be focused on broad areas, such as infrastructure, applications, and middleware. Also, there will be bursts of 10-20% dedicated effort from IT service owners when we are doing Level 3 Decompositions for specific services in 4-6 week long windows.

Office of the DCIO

As described above, the Office of the DCIO for Research IT will provide coordination for research IT activities across the campus—both within Technology Services and within various units. In some cases it will provide direct oversight for initiatives (e.g. Research IT Portal, Research User Services, Illinois Campus Cluster Program, ScienceDMZ, and Cyberinfrastructure Master Plan, etc.). In other cases it will coordinate and foster coherence among these activities in conjunction with other research IT efforts within campus units (e.g. University Library, NCSA, IGB, Beckman Institute, Technology Services, LAS, Engineering, etc.), leveraging the substantial experience and expertise of the participating units to bring additional value to the entire campus. The Office of the DCIO for Research IT will also be a resource to faculty, research professionals, students, technology professionals, and units in furthering the research IT needs of the campus. It will serve as a trusted hub of information for finding services, resources, partners, and collaborators: a function facilitated by its information gathering and dissemination activities. Further, this office will provide a means of effectively and proactively coordinating efforts to

tap into state, federal and other sources of funds to move our campus forward in providing a unified and coherent digital research ecosystem.

The Office of the DCIO for Research IT will also provide support functions for research IT initiatives. This will include business and technical processes along with developing the financial models and financial management to support these initiatives. This will be done in collaboration with the Technology Services Business and Finance office and be complemented by administrative, project management, and human resource management support. The Office of the DCIO for Research IT will establish an advisory committee to help guide the development of resources, services, and support; it is anticipated that this will leverage other such advisory activities of the campus.

In order to provide these support services to the research IT initiatives, we aggregate the needs with respect to project management (3.70 FTE initially and growing to 3.90 FTE), program management (2.05 FTE), business analyst support (1.45 FTE initially and growing to 2.20 FTE), human resources support (1.50 FTE), administrative support (1.40 FTE initially and growing to 2.15 FTE). Finally, we account for staff equipment ($2,500 per FTE), and travel and staff development ($3,000 per FTE annually) on a per FTE basis and, based on an estimate from the Technology Services Project Management Office, estimate 3.0 FTE for portfolio management.

Budget Discussion

A composite budget for the activities discussed in this document is provided in the table below. In order to illustrate leveraged efforts and commitments that exist, those commitments are shown in the table along with the request for resources necessary to support the initiatives described. The combination of funding profiles and both short term and long term commitments results in a startup estimate of approximately 89.45 FTE with an associated budget of $10.0M (including both FTE and non-FTE costs) of which approximately 21.25 FTE and a total of $2.36M (23.5%) are already committed. The ongoing estimate is approximately 120.75 FTE with an associated budget of $11.2M (including both FTE and non-FTE costs) of which approximately 28.00 FTE and a total of $2.38M (21.2%) are already committed. This leaves a net need for startup of approximately 68.20 FTE with an associated budget of $7.67M and an ongoing need of approximately 92.75 FTE with an associated budget of $8.87M.

DCIO Research IT Budget Estimates	Startup						On-Going					
	Committed			Request			Committed			Request		
	FTE	FTE $	Non-FTE $	FTE	FTE $	Non-FTE $	FTE	FTE $	Non- FTE $	FTE	FTE $	Non-FTE $
Research IT Support												
Training	3.50	$ 262,500	$ -	7.20	$ 540,000	$ 10,000	3.50	$ 262,500	$ -	8.45	$ 633,750	$ 10,000
Communications and Marketing	1.50	$ 112,500	$ -	2.30	$ 172,500	$ -	1.00	$ 75,000	$ -	3.75	$ 281,250	$ -
Research User Services	0.00	$ -	$ -	5.20	$ 390,000	$ -	1.75	$ 131,250	$ -	12.70	$ 952,500	$ -
Research Applications Support and Research Software Development	0.00	$ -	$ -	7.50	$ 562,500	$ 91,000	0.00	$ -	$ -	10.00	$ 750,000	$ 33,000
Data Visualization and Analysis	0.00	$ -	$ -	5.00	$ 375,000	$ 50,000	0.00	$ -	$ -	6.00	$ 450,000	$ 50,000
Research IT Portal	0.00	$ -	$ -	9.20	$ 652,500	$ 30,000	0.00	$ -	$ -	10.10	$ 757,500	$ 30,000
Total	**5.00**	**$ 375,000**	**$ -**	**36.40**	**$2,692,500**	**$ 181,000**	**6.25**	**$ 468,750**	**$ -**	**51.00**	**$3,825,000**	**$ 123,000**
Research Computing Services												
CARNE (ScienceDMZ)	0.25	$ 18,750	$ -	4.00	$ 300,000	$ 350,000	0.25	$ 18,750	$ -	4.00	$ 300,000	$ 300,000
Illinois Campus Cluster	0.00	$ -	$ -	0.40	$ 30,000	$ 960,000	4.25	$ 318,750	$ -	4.40	$ 330,000	$ 662,680
High Throughput Computing Service	0.00	$ -	$ -	2.00	$ 150,000	$ 25,000	0.00	$ -	$ -	1.00	$ 75,000	$ 20,000
VMs, containers, etc. for research support	0.00	$ -	$ -	3.50	$ 262,500	$ 30,000	0.00	$ -	$ -	3.50	$ 262,500	$ 30,000
Cloud Computing Service for Research	4.00	$ 300,000	$ -	1.40	$ 105,000	$ 25,000	2.00	$ 150,000	$ -	1.40	$ 105,000	$ 25,000
Total	**4.25**	**$ 318,750**	**$ -**	**11.30**	**$ 847,500**	**$1,390,000**	**6.50**	**$ 487,500**	**$ -**	**14.30**	**$1,072,500**	**$1,037,680**
Research Data Services												
Research Data Service	4.00	$ 385,000	$536,000	0.00	$ -	$ -	4.00	$ 385,000	$ 65,000	0.00	$ -	$ -
Sensitive Data Service	2.50	$ 187,500	$ -	6.00	$ 450,000	$ 500,000	5.00	$ 375,000	$ -	12.00	$ 900,000	$ 333,333
Total	**6.50**	**$ 572,500**	**$ 536,000**	**6.00**	**$ 450,000**	**$ 500,000**	**9.00**	**$ 760,000**	**$ 65,000**	**12.00**	**$ 900,000**	**$ 333,333**
Research IT Strategy & Planning												
Needs Collection and Analysis	0.00	$ -	$ -	3.20	$ 240,000	$ 40,000	0.00	$ -	$ -	3.45	$ 258,750	$ 20,000
Cyberinfrastructure Master Plan	0.00	$ -	$ -	0.50	$ 37,500	$ 100,000	0.00	$ -	$ -	0.25	$ 18,750	$ 20,000
UIUC IT Architecture	1.50	$ 112,500	$ -	1.30	$ 97,500	$ -	1.50	$ 112,500	$ -	1.30	$ 97,500	$ -
Total	**1.50**	**$ 112,500**	**$ -**	**5.00**	**$ 375,000**	**$ 140,000**	**1.50**	**$ 112,500**	**$ -**	**5.00**	**$ 375,000**	**$ 40,000**
Office of the DCIO												
Administrative Assistant	0.40	$ 30,000	$ -	1.00	$ 75,000	$ -	0.40	$ 30,000	$ -	1.75	$ 131,250	$ 10,000
Human Resources	0.00	$ -	$ -	1.50	$ 112,500	$ 10,000	0.00	$ -	$ -	1.50	$ 112,500	$ 10,000
Travel and Staff Development	0.00	$ -	$ -	0.00	$ -	$ 204,600	0.00	$ -	$ -	0.00	$ -	$ 278,250
Equipment	0.00	$ -	$ -	0.00	$ -	$ 170,500	0.00	$ -	$ -	0.00	$ -	$ 77,292
Business Analyst	0.25	$ 18,750	$ -	1.20	$ 90,000	$ -	0.50	$ 37,500	$ -	1.70	$ 127,500	$ -
Project Manager	1.10	$ 82,500	$ -	2.60	$ 195,000	$ -	1.60	$ 120,000	$ -	2.30	$ 172,500	$ -
Portfolio Manager	0.00	$ -	$ -	3.00	$ 225,000	$ -	0.00	$ -	$ -	3.00	$ 225,000	$ -
Program Manager	1.85	$ 216,000	$ -	0.20	$ 15,000	$ -	1.85	$ 201,000	$ -	0.20	$ 15,000	$ -
Other	0.40	$ 100,000	$ -	0.00	$ -	$ -	0.40	$ 100,000	$ -	0.00	$ -	$ -
Total	**4.00**	**$ 447,250**	**$ -**	**9.50**	**$ 712,500**	**$ 385,100**	**4.75**	**$ 488,500**	**$ -**	**10.45**	**$ 783,750**	**$ 375,542**
Total Grand Total	**21.25**	**$1,826,000**	**$ 536,000**	**68.20**	**$5,077,500**	**$2,596,100**	**28.00**	**$2,317,250**	**$ 65,000**	**92.75**	**$6,956,250**	**$1,909,555**
Sum of FTE and non-FTE $:			***$2,362,000***			***$7,673,600***			***$2,382,250***			***$8,865,805***

Appendix I: Year of Cyberinfrastructure Final Report

For completeness and due to the heavy referencing of the Year of Cyberinfrastructure Final Report, it is included here for easy reference.

[NOTE: The appendices for the proposal are provided as a separate document, "DCIO-ResearchIT_Proposal-v2.0-Appendicies.pdf"]

Appendix II: NewSoft Documents

The planning process has utilized a standardized process utilized in Technology Services for introducing new services. While the NewSoft process and document is mildly problematic in some of the cases here, it has supported a consistent process for representing the needs and solutions. These document constitute the "homework" done in preparation for the development of this document and are provided here for further background and detail for the interested reader.

[NOTE: The appendices for the proposal are provided as a separate document, "DCIO-ResearchIT_Proposal-v2.0-Appendicies.pdf"]